Quantum speedup of classical mixing processes*

Peter C. Richter†

Abstract

It is known that repeated measurements performed at uniformly random times enable the continuous-time quantum walk on a finite set $\mathcal{S}$ (using a stochastic transition matrix P as the time-independent Hamiltonian) to sample almost uniformly from $\mathcal{S}$ provided that P does. Here we show that the same phenomenon holds for other (discrete-time) walk variants and more general measurements types, then focus our attention on two questions: How are these repeatedly-measured walks related to the decohering quantum walks proposed by Kendon/Tregenna and Alagic/Russell? And, when do they yield a speedup over their classical counterparts?

We answer the first question with a proof that the two quantum walk models are essentially equivalent (in that they sample almost uniformly from $\mathcal{S}$ with nearly the same efficiency) by relating the spectral gaps of the Markov chains describing their action on $\mathcal{S}$. We answer the second question (in part) by showing that these quantum walks sample almost uniformly from the torus $\mathbb{Z}_n^d$ in time $O(n \log \epsilon^{-1})$. This represents a quadratic speedup over classical and for $d = 1$ confirms a conjecture of Kendon and Tregenna based on numerical experiments.

1 Introduction

Grover's algorithm [13] provides a quadratic quantum speedup (over the best classical algorithm) for the black-box problem of *bounded-error search*: return a marked element from a finite set $\mathcal{S}$ with success probability 2/3, or (in the *decision* version of the problem) simply report whether or not $\mathcal{S}$ contains a marked element with success probability 2/3. Szegedy's algorithm [24, 25] preserves this quadratic speedup in situations which require us to move through $\mathcal{S}$ using a symmetric Markov chain (or unitary matrix, in the quantum case) with prescribed locality. Such *quantum walks* have been used to develop faster quantum algorithms for several fundamental problems, including element distinctness (Ambainis [4]), matrix product verification (Burhman and Spalek [6]), triangle finding (Magniez, Santha, and Szegedy [20]), subset finding (Childs and Eisenberg [8]), and group commutativity testing (Magniez and Nayak [19]).

This paper concerns the use of quantum walks in yielding a speedup for another basic problem, that of *almost uniform sampling*: return an element from within ϵ total variation distance of the uniform distribution over a finite set $\mathcal{S}$. As in the setup above, we require the algorithm (whether classical or quantum) to respect a particular locality structure on $\mathcal{S}$. The classical *Markov chain Monte Carlo* solution to this problem is a key component underlying most approximation algorithms for #P-complete problems [14]. The possibility of obtaining a quantum speedup for this problem

*This material is based upon work supported by the National Science Foundation under Grant No. 0523866. Part of this research was done while the author was visiting LRI, Université Paris-Sud, Orsay, France.

†Department of Computer Science, Rutgers University, Piscataway, NJ 08854. `richterp@cs.rutgers.edu`

has been considered by Aharonov et al. [1], Nayak et al. [22, 5], Moore and Russell [21], Kendon and Tregenna [17], Gerhardt and Watrous [12], Alagic and Russell [2], and Richter [23].

We make three contributions in this area: First, we give a simple argument that any reasonable unitary quantum walk transition rule (including those most familiar from the literature) can be used for almost uniform sampling provided that measurements are performed with some randomness and frequency. This property of global convergence to the uniform classical state applies more generally to any time-independent quantum dynamics subject to time-averaged measurements or decoherence. This generalizes Theorem 3.5 of [23].

Second, we show (in the continuous-time setting) that two of the three common measurement rules for quantum walks (the *uniform*, or Cesaro-averaging, rule of Aharonov et al. [1, 21, 12, 23] and the *memoryless*, or Bernoulli/Poisson-averaging, rule of Kendon et al. [17, 2]) are equivalent in the following sense: if the almost uniform sampling problem can be solved by a quantum walk using one of the two measurement rules, then it can be solved in nearly the same time using the other measurement rule.[1] This resolves an open question of [23]. The proof applies more generally to a particular game involving classical Markov chains (not necessarily describing quantum phenomena) and may be of independent interest.

Third, we prove upper bounds on the time complexity of sampling almost uniformly using continuous- and discrete-time quantum walks with uniform, memoryless, or instantaneous measurement rules. In particular, we show that such quantum walks can be used to sample from the torus $\mathbb{Z}_n^d$ in time $O(n \log \epsilon^{-1})$. This represents a quadratic speedup over the classical random walk behavior and resolves conjectures of Kendon and Tregenna [17] (for $d = 1$) and Richter [23].

In an appendix, we show that the continuous- and discrete-time quantum walks on the hypercube $\mathbb{Z}_2^n$ using time-averaged (either uniform or memoryless) measurements with the natural parameter $T = O(n)$ (to allow the quantum wavefunction just enough time to traverse the hypercube before collapsing from measurement) sample almost uniformly in time $O(n^{3/2} \log \epsilon^{-1})$. This proves a weak form of the amplification conjecture for $\mathbb{Z}_2^n$ in [23]. The result is derived in two ways (i.e., the continuous- and discrete-time proofs differ considerably) and if tight would represent a *slowdown* over both the classical random walk (which succeeds in time $O(n \log n)$) and the quantum walk with instantaneous measurement (which succeeds in time $O(n)$ [21]).

It remains a largely open problem to characterize the class of Markov chains that permit a quantum speedup.

2 Preliminaries

2.1 Markov chains

The classical Markov chain Monte Carlo method for almost uniform sampling from a set $\mathcal{S}$ of size N works as follows: Run a Markov chain (i.e., iterate a stochastic matrix) P on $\mathcal{S}$ until it is guaranteed to output a state ϵ-close to the uniform distribution u over $\mathcal{S}$ in *total variation* distance $\frac{1}{2}|| \cdot ||_1$, regardless of the initial state. We require P to be (a) symmetric so that u is a *stationary distribution* (i.e., $Pu = u$), (b) *irreducible* (strongly connected) so that u is the *unique* stationary distribution, and (c) *aperiodic* (non-bipartite) to ensure global convergence to u from any initial state. An irreducible, aperiodic Markov chain is called *ergodic*. The *mixing time*

[1] The *instantanteous* rule [22, 5, 21, 2] is incomparable to the other two in this way.

$\tau(\epsilon) := \min\{T : \frac{1}{2}||P^t - u1^\dagger||_1 \leq \epsilon \, \forall t \geq T\}$ is the minimum number of iterations of P required to output a state ϵ-close to uniform.[2]

The following well-known inequality relates the mixing time of a symmetric (or more generally, reversible) ergodic Markov chain to its *spectral gap* $\delta := 1 - \lambda$, where $\lambda := ||P|_{u^\perp}||_2$:

Theorem 2.1 (Diaconis, Strook [10]; Aldous [3]) *Let P be symmetric and ergodic. Then its mixing time satisfies* $\frac{1}{2}\lambda\delta^{-1}\ln(2\epsilon)^{-1} \leq \tau(\epsilon) \leq \delta^{-1}(\ln N + \ln \epsilon^{-1})$.

The $\Theta(\log N)$ slack between upper and lower bounds is necessary in general, but tight asymptotic estimates are known for some Markov chains. For example, the simple random walks on the hypercube $\mathbb{Z}_2^n$ and the torus $\mathbb{Z}_n^d$ mix in time $\Theta(n \log n)$ and $\Theta(n^2)$, respectively.

We will also use the *maximum pairwise column distance* $\bar{d}(P) := \max_{x,x'} \frac{1}{2}||P(\cdot, x) - P(\cdot, x')||_1$ to estimate the mixing time. It is related to the total variation distance from the uniform (or more generally, stationary) distribution by the inequality:

$$\frac{1}{2}||P - u1^\dagger||_1 \leq \bar{d}(P) \leq ||P - u1^\dagger||_1 \tag{1}$$

The following propositions (see [23]) can be used to estimate the mixing time of P given a common lower bound on most of the entries in each column:

Proposition 2.2 *If $\bar{d}(P) \leq \alpha$, then $\tau(\epsilon) \leq \lceil \log_{1/\alpha} \epsilon^{-1} \rceil$.*

Proposition 2.3 *If at least βN entries in each column of P are bounded below by γ/N, where $\beta > \frac{1}{2}$ and $\gamma > 0$, then $\bar{d}(P) \leq 1 - \gamma(1 - 2(1 - \beta))$.*

2.2 Quantum walks

A *quantum walk* on the set $\mathcal{S}$ takes place on the Hilbert space spanned by the elements of $\mathcal{S}$ (tensored with an auxiliary Hilbert space, possibly) and is specified by two components: a transition rule and a measurement rule. A *transition rule* is a Hamiltonian H (in the continuous-time setting) or unitary matrix U (in the discrete-time setting) whose nonzero entries respect the locality structure imposed by the problem constraints.[3] A *measurement rule* is a family ω_T of probability density functions on $[0, \infty)$ (in the continuous-time setting) or probability mass functions on $[0, \infty) \cap \mathbb{Z}$ characterizing the (random) time at which a total measurement on the Hilbert space is performed, collapsing the wavefunction.[4] For a quantum walk $\{H, \omega_T\}$ (in continuous time) or $\{U, \omega_T\}$ (in discrete time), let us define the stochastic matrix P_t by $|\langle y|e^{-iHt}|x\rangle|^2$ or $|\langle y|U|x\rangle|^2$, respectively. The effect of the quantum walk is to implement the stochastic matrix $E_{\omega_T}[P_t] := \int \omega_T(t) \, dt \, P_t$.

The following transition rules have been proposed for quantum walks: Given a symmetric Markov chain P, the *continuous-time* walk [11, 9] is given by $U_t = e^{-iPt}$ (i.e., time-independent dynamics with Hamiltonian $H = P$). The discrete-time walk is slightly more complex and comes in several variants; we shall focus on the main three: the Hadamard, DFT, and Grover/Szegedy walks. All variants can be viewed as taking place on the product Hilbert space $\{|x\rangle|y\rangle : x, y \in \mathcal{S}\}$ (indeed, in the subspace spanned by all $|x\rangle|y\rangle$ such that x to y is an arc, or allowed transition, in

[2] We use $1^\dagger$ to denote the N-dimensional row vector of ones; $\dagger$ is the conjugate transpose.

[3] In the discrete-time setting, this construction is only possible if we use an auxiliary "coin" space to label the allowed transitions from an element of $\mathcal{S}$, so we can record movement between elements of $\mathcal{S}$ in a reversible manner.

[4] We could in principle consider more general types of measurements, but this suffices for our purposes.

the locality structure) and are given by $U_t = U^t$, where $U := RS$ is the product of a *shift operator* $S := \sum_{x,y\in\mathcal{S}} |x,y\rangle \mapsto |y,x\rangle$ and a *pivot operator* $R := \sum_{x\in\mathcal{S}} \Pi_x \otimes C_x$. Here Π_x is the projector onto $|x\rangle$, and C_x is the *coin operator* that determines the walk variant. The *Hadamard* walk [18] on $\mathcal{S} = \mathbb{Z}_m^n$ or $\mathcal{S} = \mathbb{Z}^n$ is obtained by setting $C_x := \bigotimes_{j=1}^n H'_{x_j}$, where H'_{x_j} is the Hadamard transform H on the subspace $\{|x_j-1\rangle, |x_j+1\rangle\}$ and the identity elsewhere. The *DFT* and *Grover* walks [15] on the vertices $\mathcal{S}$ of an undirected graph are obtained by letting C_x be the discrete Fourier transform and Grover diffusion operator, respectively, on the subspace spanned by x and its neighbors and the identity elsewhere.[5] The *Szegedy* walk [24, 25] is a generalization of the Grover walk; it "quantizes" an arbitrary Markov chain by letting C_x be a reflection through the vector $\sum_{y\in\mathcal{S}} \sqrt{P(y,x)}|y\rangle$ rather than $\frac{1}{\sqrt{N}} \sum_{y\in\mathcal{S}} |y\rangle$.

The following quantum walk measurement rules have been proposed: The *instantaneous* rule [22, 5] is simply the point distribution $\xi_T(t) := \delta(t-T)$, where δ is the delta function. The *uniform* rule [1] is the uniform distribution $\bar{\mu}_T := \frac{1}{T}\chi_{[0,T]}$ (in continuous time) or $\bar{\nu}_T := \frac{1}{T}\chi_{[0..T-1]}$ (in discrete time), where χ is the characteristic function. The *memoryless* rule is the exponential distribution $\tilde{\mu}_T(t) := \frac{1}{T}\exp(-t/T)$ (in continuous time) or the geometric distribution $\tilde{\nu}_T(t) := \frac{1}{T}(1-\frac{1}{T})^t$ (in discrete time). It describes the interarrival time between measurements in a Poisson process with measurements occurring at rate $\lambda = 1/T$ (in continuous time) or a Bernoulli process with measurements occurring with probability $p = 1/T$ at each timestep (in discrete time). These processes coincide with the decoherence (measurement) models of Alagic and Russell [2] and Kendon and Tregenna [17], respectively. Their decoherence models are parametrized by the decoherence rate λ or probability p and the walk length $n \gg 1/p$; thus, their quantum walks undergo multiple measurements (of memoryless type). The ability of these multiple measurements to enhance mixing was first demonstrated in numerical experiments by Kendon and Tregenna [17]. More recently, mathematical justification for the ability of multiple *uniform* measurements to enhance mixing has been given by Richter [23].

3 The effects of measurement on quantum walk dynamics

3.1 Single versus multiple measurements

As we have already noted, the effect of a quantum walk with transition rule H (in continuous time) or U (in discrete time) and measurement rule ω_T is to implement the stochastic matrix $E_{\omega_T}[P_t]$, where P_t is the matrix $|\langle y|e^{-iHt}|x\rangle|^2$ (in continuous time) or $|\langle y|U^t|x\rangle|^2$ (in discrete time). Thus, the effect of running this walk (or equivalently, repeating the measurement rule) T' times is to implement the stochastic matrix $(E_{\omega_T}[P_t])^{T'}$. The next lemma and theorem (generalizing Theorem 3.4 of [1] and Theorem 3.5 of [23]) describe the asymptotic behavior of $E_{\omega_T}[P_t]$ and $(E_{\omega_T}[P_t])^{T'}$ in the limits $T \to \infty$ and $T' \to \infty$, respectively. Although stated explicitly for quantum walks, they may be of more general interest in that they simply describe the long-term behavior of time-independent quantum dynamics on a finite-dimensional Hilbert space subjected to random destructive measurements.

Lemma 3.1 (Single-measurement dynamics) *Let $\{H, \omega_T\}$ (in continuous time) or $\{U, \omega_T\}$ (in discrete time) be a quantum walk. If $E_{\omega_T}[e^{i\theta t}] \to 0$ as $T \to \infty$ for any $\theta \neq 0$, then*

$$E_{\omega_T}[P_t] \to \Pi \textit{ as } T \to \infty \tag{2}$$

[5]Note that the Hadamard and DFT walks are identical in one dimension (i.e., on the line $\mathbb{Z}$ or cycle $\mathbb{Z}_N$).

where Π *is the stochastic matrix given by* $\Pi(f,e) := \sum_j \left|\sum_{k\in C_j}\langle f|\phi_k\rangle\langle\phi_k|e\rangle\right|^2$, $\{\lambda_k, |\phi_k\rangle\}$ *is the spectrum of* H *(or* U*), and* $\{C_j\}$ *is the partition of these indices* k *obtained by grouping together the* k *with identical* λ_k*.*

Proof: Decomposing the walk along spectral components gives us

$$E_{\omega_T}[P_t](f,e) = E_{\omega_T}[|\sum_k \langle f|\phi_k\rangle\langle\phi_k|e\rangle e^{-i\theta_k t}|^2] \tag{3}$$

where $\theta_k := \lambda_k$ (in continuous time) or $e^{-i\theta_k} := \lambda_k$ (in discrete time). Writing $|\cdot|^2$ as a product of complex conjugates, we obtain:

$$\begin{aligned} E_{\omega_T}[P_t](f,e) &= E_{\omega_T}[(\sum_k \langle f|\phi_k\rangle\langle\phi_k|e\rangle)(\sum_l \langle\phi_l|f\rangle\langle e|\phi_l\rangle)e^{-i(\theta_k-\theta_l)t}] && (4)\\ &= (\sum_k \langle f|\phi_k\rangle\langle\phi_k|e\rangle)(\sum_l \langle\phi_l|f\rangle\langle e|\phi_l\rangle)E_{\omega_T}[e^{i(\theta_l-\theta_k)t}] && (5)\end{aligned}$$

Now by assumption, $E_{\omega_T}[e^{i(\theta_l-\theta_k)t}]$ vanishes as $T\to\infty$ for all $\theta_l \neq \theta_k$, so we have

$$E_{\omega_T}[P_t](f,e) \to \sum_k \langle f|\phi_k\rangle\langle\phi_k|e\rangle(\sum_{l:\theta_l=\theta_k}\langle\phi_k|f\rangle\langle e|\phi_k\rangle) = \sum_j |\sum_{k\in C_j}\langle f|\phi_k\rangle\langle\phi_k|e\rangle|^2 = \Pi(f,e) \tag{6}$$

in the limit $T\to\infty$. ■

Theorem 3.2 (Multiple-measurement dynamics) *Let* $\{H,\omega_T\}$ *(in continuous time) or* $\{U,\omega_T\}$ *(in discrete time) be a quantum walk. If the underlying graph of* H *(or* U*) is strongly connected and non-bipartite, then for* T *sufficiently large:*

$$(E_{\omega_T}[P_t])^{T'} \to u1^\dagger \text{ as } T'\to\infty \tag{7}$$

Proof: We need to show that for T sufficiently large, the Markov chain $E_{\omega_T}[P_t]$ is ergodic with uniform stationary distribution.

That the uniform distribution is stationary is clear: Each of the P_t already has uniform stationary distribution, since the uniform classical state is invariant under unitary quantum operations and under total measurement of the system. Thus, any probabilistic combination of them also has uniform stationary distribution.

To show that $E_{\omega_T}[P_t]$ is ergodic for all sufficiently large T, it is sufficient (by Lemma 3.1) to prove that Π is ergodic.[6] In the continuous-time setting, note (using a Taylor series expansion) that if $\langle f|H|e\rangle \neq 0$, then $\langle f|e^{-iHt}|e\rangle \neq 0$ for sufficiently small $t>0$, and

$$\begin{aligned} \langle f|e^{-iHt}|e\rangle \neq 0 &\Rightarrow \sum_k \langle f|\phi_k\rangle\langle\phi_k|e\rangle e^{-i\lambda_k t} \neq 0 && (8)\\ &\Rightarrow \sum_j e^{-i\lambda_{k:k\in C_j}t}\sum_{k\in C_j}\langle f|\phi_k\rangle\langle\phi_k|e\rangle \neq 0 && (9)\\ &\Rightarrow \exists j:\ |\sum_{k\in C_j}\langle f|\phi_k\rangle\langle\phi_k|e\rangle|^2 > 0 && (10)\end{aligned}$$

[6] The latter implies the former because the ergodic matrices form an open subset of the set of stochastic matrices.

so $\Pi(f, e) > 0$. Analogously, in the discrete-time setting we have

$$\langle f|U|e\rangle \neq 0 \quad \Rightarrow \quad \sum_k \langle f|\phi_k\rangle\langle\phi_k|e\rangle\lambda_k^t \neq 0 \tag{11}$$

$$\Rightarrow \quad \sum_j \lambda_{k:k\in C_j}^t \sum_{k\in C_j} \langle f|\phi_k\rangle\langle\phi_k|e\rangle \neq 0 \tag{12}$$

$$\Rightarrow \quad \exists j : |\sum_{k\in C_j} \langle f|\phi_k\rangle\langle\phi_k|e\rangle|^2 > 0 \tag{13}$$

so again $\Pi(f, e) > 0$. Thus, Π has (at least) the same nonzero entries that H (or U) has. In particular, if the graph underlying H (or U) is strongly connected and non-bipartite, then so is that of Π; hence Π is ergodic. ■

It is easy to check that any quantum walk using a standard transition rule (Hadamard, DFT, or Grover/Szegedy) and either a uniform or memoryless measurement rule (but not the instantaneous rule) satisfies the conditions of the above lemma and theorem.[7]

The question of which classical random walks can be "quantized" to yield a quantum walk with mixing time $T \cdot T'$ significantly less than the classical mixing time remains largely open.

3.2 Memoryless versus uniform measurements

It was asked in [23] whether memoryless measurements [17, 2] and uniform measurements [1, 23] solve the almost uniform sampling problem with essentially the same efficiency. We answer this question in the affirmative.[8]

Heretofore, let $\bar{P}_T := E_{\bar{\mu}_T}[P_t]$, $\tilde{P}_T := E_{\tilde{\mu}_T}[P_t]$, $\bar{Q}_T := E_{\bar{\nu}_T}[P_t]$, and $\tilde{Q}_T := E_{\tilde{\nu}_T}[P_t]$. Also let $\bar{\delta}_T := 1 - ||\bar{P}_T|_{u^\perp}||_2$, $\tilde{\delta}_T := 1 - ||\tilde{P}_T|_{u^\perp}||_2$, $\bar{\zeta}_T := 1 - ||\bar{Q}_T|_{u^\perp}||_2$, and $\tilde{\zeta}_T := 1 - ||\tilde{Q}_T|_{u^\perp}||_2$.

Lemma 3.3 (Spectral gap inequalities) *Let $\bar{\delta}_T$, $\tilde{\delta}_T$, $\bar{\zeta}_T$, and $\tilde{\zeta}_T$ be defined as above. Then for any $k \geq 1$ we have the inequalities*

$$e^{-1}\bar{\delta}_T \leq \tilde{\delta}_T \leq k(1 - e^{-k}) \cdot \bar{\delta}_{kT} + 2e^{-k} \tag{14}$$

and

$$4^{-1}\bar{\zeta}_T \leq \tilde{\zeta}_T \leq k(1 - e^{-k}) \cdot \bar{\zeta}_{kT} + 2e^{-k} \tag{15}$$

Proof: We will prove the continuous-time version; the discrete-time version is nearly identical.

Suppose we want to simulate $\bar{P}_T$ by $\tilde{P}_T$. Scaling the distribution $\bar{\mu}_T$ by $\alpha := 1/e$ allows us to "fit it inside" the distribution $\tilde{\mu}_T$ (i.e., $e^{-1}\bar{\mu}_T \leq \tilde{\mu}_T$ pointwise), so we can express $\tilde{\mu}_T$ as the probabilistic combination $\alpha\bar{\mu}_T + (1-\alpha)\nu$ for some distribution ν, so that

$$\tilde{P}_T = E_{\tilde{\mu}_T}[P_t] = \alpha E_{\bar{\mu}_T}[P_t] + (1-\alpha)E_\nu[P_t] = \alpha\bar{P}_T + (1-\alpha)Q \tag{16}$$

where Q is stochastic with uniform stationary distribution. It follows that

$$||\tilde{P}_T|_{u^\perp}||_2 \leq 1/e||\bar{P}_T|_{u^\perp}||_2 + (1 - 1/e)||Q|_{u^\perp}||_2 \tag{17}$$

[7] Except the Grover/Szegedy walk on $\mathbb{Z}_N$, which tends to the uniform distribution on one of the directed cycles.

[8] One might argue that the walk of Kendon and Tregenna [17] uses the walk length rather than the number of measurements as a parameter, however when the number of measurements is fixed to a large value the walk length is already determined (asymptotically) using a Chernoff bound.

which implies that $\tilde{\delta}_T \geq 1/e \cdot \bar{\delta}_T$ since $||Q|_{u^\perp}||_2 \leq 1$.

Suppose we want to simulate $\tilde{P}_T$ by $\bar{P}_{kT}$. Then the basic approach is the same, but since the support of $\tilde{\mu}_T$ is not compact we have to be careful. Scaling the distribution $\tilde{\mu}_T$ by $\beta := 1/k$ allows us to fit it inside the distribution $\bar{\mu}_{kT}$ up to the point $t = kT$, and the probability mass in $\tilde{\mu}_T$ past $t = kT$ is only $\Pr_{\tilde{\mu}_T}[t > kT] = e^{-k}$. So we can write

$$\tilde{\mu}_T = (1 - e^{-k}) \cdot \tilde{\mu}_T^{head} + e^{-k} \cdot \tilde{\mu}_T^{tail} \tag{18}$$

where $\tilde{\mu}_T^{head}$ and $\tilde{\mu}_T^{tail}$ are the conditional distributions of $\tilde{\mu}_T$ such that $t \leq kT$ and $t > kT$, respectively; thus,

$$\tilde{P}_T = (1 - e^{-k}) \cdot \tilde{P}_T^{head} + e^{-k} \cdot \tilde{P}_T^{tail} \tag{19}$$

where $\tilde{P}_T^{head}$ and $\tilde{P}_T^{tail}$ are the expectations of P_t w.r.t. $\tilde{\mu}_T^{head}$ and $\tilde{\mu}_T^{tail}$, respectively. Since we can fit $\tilde{\mu}_T^{head}$ inside $\bar{\mu}_{kT}$ if we scale it by $1/k$, we can write

$$\bar{P}_{kT} = \frac{1}{k}\tilde{P}_T^{head} + (1 - \frac{1}{k})Q \tag{20}$$

where Q is stochastic with uniform stationary distribution. The above equations yield:

$$\begin{aligned} \bar{P}_{kT} &= \frac{1}{k(1-e^{-k})}(\tilde{P}_T - e^{-k}\tilde{P}_T^{tail}) + (1 - \frac{1}{k})Q \\ &= \frac{1}{k(1-e^{-k})}\tilde{P}_T - \frac{e^{-k}}{k(1-e^{-k})}\tilde{P}_T^{tail} + (1 - \frac{1}{k})Q \end{aligned} \tag{21}$$

From the triangle inequality, we obtain

$$||\bar{P}_{kT}|_{u^\perp}||_2 \leq \frac{1}{k(1-e^{-k})}||\tilde{P}_T|_{u^\perp}||_2 + \frac{e^{-k}}{k(1-e^{-k})}||\tilde{P}_T^{tail}|_{u^\perp}||_2 + (1 - \frac{1}{k})||Q|_{u^\perp}||_2 \tag{22}$$

and, rearranging terms and simplifying:

$$\frac{1}{k(1-e^{-k})}(1 - ||\tilde{P}_T|_{u^\perp}||_2) - \frac{2e^{-k}}{k(1-e^{-k})} \leq 1 - ||\bar{P}_{kT}|_{u^\perp}||_2 \tag{23}$$

■

In the following theorem, we simplify the expressions by taking the target distance ϵ from uniform to be any small positive constant. We do so without loss of generality, since the dependence on ϵ always appears as an $O(\log \epsilon^{-1})$ factor [23].

Theorem 3.4 (Equivalence of measurement rules) *Fix a quantum walk transition rule on $\mathcal{S}$ for which P_t is symmetric. Then: (a) If T' uniform measurements $\bar{\mu}_T$ (or $\bar{\nu}_T$) are sufficient to sample almost uniformly from $\mathcal{S}$, then $T' \cdot O(\log N)$ memoryless measurements $\tilde{\mu}_T$ (or $\tilde{\nu}_T$) are also sufficient; (b) If T' memoryless measurements $\tilde{\mu}_T$ (or $\tilde{\nu}_T$) are sufficient to sample almost uniformly from $\mathcal{S}$, then $T' \cdot O(\log T' \log N)$ uniform measurements $\bar{\mu}_{T \cdot O(\log T')}$ (or $\bar{\nu}_{T \cdot O(\log T')}$) are also sufficient.*

Proof: We will prove the continuous-time version; the discrete-time version is nearly identical.

To see (a), note that our assumption implies that $\bar{P}_T$ mixes in time T'. Therefore, $\bar{\delta}_T = \Omega(\lambda/T')$ by Theorem 2.1, where $\lambda = 1 - \bar{\delta}_T$. We will assume that λ is bounded away from 0; if it were not, then $\bar{\delta}_T$ would be, and the quantum walk would be unnecessary because the classical walk would already mix optimally fast. Then $\bar{\delta}_T = \Omega(1/T')$, and from Lemma 3.3 it follows that $\tilde{\delta}_T = \Omega(1/T')$. Applying Theorem 2.1 again, we obtain for $\tilde{P}_T$ a mixing time of $O(T' \log N)$.

The proof of (b) is almost as straightforward. Our assumption implies that $\tilde{P}_T$ mixes in time T', so $\tilde{\delta}_T = \Omega(1/T')$ by Theorem 2.1. Set k to be the smallest integer for which $\tilde{\delta}_T \geq 3e^{-k}$; in particular, $k = \Theta(\log \tilde{\delta}_T^{-1}) = O(\log T')$. By Lemma 3.3:

$$\bar{\delta}_{kT} \geq \frac{1}{k(1-e^{-k})}(\tilde{\delta}_T - 2e^{-k}) \geq \frac{1}{k(1-e^{-k})}(e^{-k}) = \Theta(\frac{\tilde{\delta}_T}{\log \tilde{\delta}_T^{-1}}) = \Theta(\frac{1}{T' \log T'}) \tag{24}$$

Applying Theorem 2.1 again, we obtain for $\bar{P}_{kT}$ a mixing time of $O(T' \log T' \log N)$. ∎

Among the standard quantum walk transition rules, only the continuous-time variant is guaranteed to produce symmetric P_t. The symmetry (or more generally, reversibility) requirement is essential because the mixing time lower bound of Theorem 2.1 fails spectacularly for some irreversible Markov chains. Nevertheless, it is likely that in most situations a similar discrete-time equivalence holds.

It should be readily apparent that the equivalence result holds for any two measurement rules with finite expectation and significant overlap for most T. We also remark that although the above lemma and theorem are stated in terms of quantum walks, the proofs indicate that they are merely statements about an abstract game involving a collection of symmetric Markov chains $\{P_t\}_{t \geq 0}$ and a T-parametrized family of probability measures $\{\omega_T\}$, where we seek to minimize the "cost function" $T \cdot T'$.

4 Quantum speedup of mixing on the torus

The instantaneous measurement rule has been shown to enable almost uniform sampling from the cycle $\mathbb{Z}_N$ in optimal time $O(N)$ (cf. [22, 5, 7]). This represents a quadratic speedup over the classical behavor of $O(N^2)$. We extend this to an $O(n)$ upper bound for the continuous-time and Hadamard walks on the torus $\mathbb{Z}_n^d$, also a quadratic speedup, and show that it holds for uniform and memoryless measurements as well. In particular, for $d = 1$ this resolves a conjecture of Kendon and Tregenna [17] based on numerical experiments.

Theorem 4.1 (Continuous-time walk on $\mathbb{Z}_n^d$) *Using either instantaneous, uniform, or memoryless measurements with parameter $T = O(n)$, the continuous-time walk samples almost uniformly from $\mathbb{Z}_n^d$ in time $O(n \log \epsilon^{-1})$.*

Proof: Set $T = \frac{n}{2}$, and consider the time interval $\mathcal{I} := ((1+\delta) \cdot (\frac{2}{3})^{1/d} \frac{n}{2}, (1-\delta) \cdot \frac{n}{2})$ inside $[0, T]$, where $\delta > 0$ is a small constant. We will show that $\bar{d}(P_t)$ is bounded below one by a positive constant for every $t \in \mathcal{I}$. Since both $\bar{\mu}_T$ and $\tilde{\mu}_T$ output a random variable t landing in the interval $\mathcal{I}$ with constant probability, it follows that both $\bar{d}(\bar{P}_T)$ and $\bar{d}(\tilde{P}_T)$ are bounded below one by a positive constant, hence $T' = O(\log \epsilon^{-1})$ suffices (by Proposition 2.2) for mixing using either uniform or memoryless measurements.

Let $|\phi_t\rangle$ and $|\psi_t\rangle$ be the wavefunctions at time t for the continuous-time walks on $\mathbb{Z}$ and $\mathbb{Z}_n^d$, respectively, starting from the origin. Then for each $\bar{y} \in \mathbb{Z}_n^d$ we have:

$$\langle \bar{y}|\psi_t\rangle = \prod_{j=1}^{d} \sum_{y_j \equiv \bar{y}_j \bmod n} \langle y_j|\phi_t\rangle \tag{25}$$

Childs [7] shows that $\langle y_j|\phi_t\rangle = (-i)^{y_j} J_{y_j}(t)$, where J_{y_j} is a Bessel function of the first kind. In particular, for $|y_j| \gg 1$ the quantity $|J_{y_j}(t)|$ is (a) exponentially small in $|y_j|$ for $t < (1-\epsilon)\cdot|y_j|$ and (b) of order $|y_j|^{-1/2}$ for $t > (1+\epsilon)\cdot|y_j|$. For every $t < (1-\delta)\cdot\frac{n}{2}$, property (a) implies that for each j the only term in the above summand that is non-negligible is the $\langle y_j|\phi_t\rangle$ with $|y_j| < \frac{n}{2}$ (call it $\hat{y}_j$ and note that $\bar{y} \leftrightarrow \hat{y}$ is a 1-1 correspondence), so we can use property (b) to conclude that

$$|\langle \bar{y}|\psi_t\rangle| \approx \prod_{j=1}^{d} |\langle \hat{y}_j|\phi_t\rangle| = \Theta((\frac{1}{\sqrt{n}})^d) \tag{26}$$

for every $t > (1+\epsilon)\cdot||\hat{y}||_\infty$. In particular, the $\frac{2}{3}n^d$ different $\bar{y}$ with $||\hat{y}||_\infty \le \frac{1}{2}(\frac{2}{3})^{1/d}n$ satisfy $|\langle \bar{y}|\psi_t\rangle| = \Omega(\frac{1}{n^{d/2}})$, and therefore $P_t(\bar{y},\bar{0}) = \Omega(\frac{1}{n^d})$, for every $t \in \mathcal{I}$. So by Proposition 2.3, $\bar{d}(P_t)$ is bounded below one by a positive constant. ■

Numerical experiments by Mackay et al. [18] indicate that the Hadamard, DFT, and Grover walks on the torus spread similarly (i.e., their standard deviations are asymptotically equivalent). We shall analyze the Hadamard walk, whose asymptotics for $d = 1$ have been computed [22, 5]. For $d > 1$, the fact that this walk is separable across the d dimensions simplifies the analysis, although a parity complication arises. To remedy it, we perturb our initial basis state to a randomized state by incrementing each coordinate (except one, if we so choose) with probability one-half before running the walk. With this simple modification, we have the following theorem:

Theorem 4.2 (Hadamard walk on $\mathbb{Z}_n^d$) *Using either instantaneous, uniform, or memoryless measurements with parameter $T = O(n)$, the Hadamard walk samples almost uniformly from $\mathbb{Z}_n^d$ in time $O(n \log \epsilon^{-1})$.*

Proof: Let $T = \frac{n}{\sqrt{2}}$ and $\mathcal{I} := ((1+\delta)\cdot(\frac{2}{3})^{1/d}\frac{n}{\sqrt{2}}, (1-\delta)\cdot\frac{n}{\sqrt{2}}) \cap \mathbb{Z}$, where $\delta > 0$ is a small constant. Say the walk starts from the state $|0,1\rangle^{\otimes n} = |\bar{0},\bar{1}\rangle$. Let $P'_t(\bar{y},\bar{0})$ be the probability for the Hadamard walk to transit from $\bar{0} \in \mathbb{Z}_n^d$ to $\bar{y} \in \mathbb{Z}_n^d$ in time t, disregarding the coin space. We will show that for every $t \in \mathcal{I}$, $P'_t(\bar{y},\bar{0}) = \Omega(1/n^d)$ for at least $\frac{2}{3}$ fraction of the $\bar{y}$ satisfying $\hat{y}_j \equiv t \bmod 2$ for all j, where $\bar{y} \leftrightarrow \hat{y}$ is the 1-1 correspondence in the proof of Theorem 4.1. However, for $d > 1$, $P'_t(\bar{y},\bar{0}) = 0$ if not all j satisfy $\hat{y}_j \equiv t \bmod 2$; in particular, P'_t is a reducible (disconnected) Markov chain. Perturbation of the initial state remedies this: Let $P''_t(\bar{y},\bar{0})$ be the transit probability when the initial state is perturbed as described above, so $P''_t(\bar{y},\bar{0}) = \frac{1}{2^{d-1}}\sum_{\bar{r}\in\{0\}\times\{0,1\}^{d-1}} P'_t(\bar{y},-\bar{r})$. Then for $t \in \mathcal{I}$, $P''_t(\bar{y},\bar{0}) = \Omega(1/n^d)$ for at least $\frac{2}{3}$ fraction of the $\bar{y}$ satisfying $\hat{y}_1 \equiv t \bmod 2$; in particular, P''_t is irreducible, but periodic (bipartite). Time-averaged (uniform or memoryless) measurement "breaks parity" (periodicity) on the remaining coordinate,[9] and it is easy to see (cf. Theorem A.2) that for at least $\frac{2}{3}$ fraction of the $\bar{y} \in \mathbb{Z}_n^d$, $E_{\bar{\nu}_T}[P''_t](\bar{y},\bar{0})$ and $E_{\tilde{\nu}_T}[P''_t](\bar{y},\bar{0})$ are bounded from below

[9] Instantaneous measurement does not, in which case we must perturb the initial state in all of the d directions.

by $\Omega(1/n^d)$. So by Proposition 2.3, $\bar{d}(P_t'')$ is bounded below one by a positive constant, and from Proposition 2.2 the theorem follows.

Let us see why for every $t \in \mathcal{I}$, $P_t'(\bar{y}, \bar{0}) = \Omega(1/n^d)$ for at least $\frac{2}{3}$ fraction of the $\bar{y}$ satisfying $\hat{y}_j \equiv t \bmod 2$ for all j. Let $|\phi_t\rangle$ and $|\psi_t\rangle$ be the wavefunctions at time t for the Hadamard walks on $\mathbb{Z}$ and $\mathbb{Z}_n^d$, respectively, starting from the states $|0, 1\rangle$ and $|\bar{0}, \bar{1}\rangle$. Then for each $\bar{y} \in \mathbb{Z}_n^d$ and $\bar{e} \in \{-1, 1\}^d$, we have:

$$\langle \bar{y}, \bar{y} + \bar{e} | \psi_t \rangle = \prod_{j=1}^{d} \sum_{y_j \equiv \bar{y}_j \bmod n} \langle y_j, y_j + e_j | \phi_t \rangle \tag{27}$$

Nayak et al. [22, 5] show that when $y_j \equiv t \bmod 2$, the quantity $|\langle y_j, y_j - 1 | \phi_t \rangle|^2 + |\langle y_j, y_j + 1 | \phi_t \rangle|^2$ is (a) exponentially small in $|y_j|$ for $t < (1 - \epsilon) \cdot \sqrt{2} |y_j|$ and (b) of order $|y_j|^{-1}$ for $(1 - \epsilon)\sqrt{2}t$ of the $y_j \in (-\frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}})$.[10] (When we do not have $y_j \equiv t \bmod 2$, the quantity is zero.) For every $t < \frac{n}{2}$, property (a) implies that for each j the only term in the above summand that is non-negligible is $\langle \hat{y}_j, \hat{y}_j + e_j | \phi_t \rangle$, so we can use property (b) to conclude that

$$|\langle \bar{y}, \bar{y} + \bar{e} | \psi_t \rangle| \approx \prod_{j=1}^{d} |\langle \hat{y}_j, \hat{y}_j + e_j | \phi_t \rangle| = \Theta((\frac{1}{\sqrt{n}})^d) \tag{28}$$

for some $\bar{e}$ for all but an arbitrarily small constant fraction of the $\bar{y}$ satisfying $\hat{y}_j \equiv t \bmod 2$ for all j and $t > (1 + \epsilon)\sqrt{2} \cdot ||\hat{y}||_\infty$, by choosing ϵ sufficiently small in relation to δ. In particular, for any $t \in \mathcal{I}$, at least $\frac{2}{3}$ fraction of the $\bar{y}$ satisfying $\hat{y}_j \equiv t \bmod 2$ for all j satisfy $|\langle \bar{y}, \bar{y} + \bar{e} | \psi_t \rangle| = \Omega(\frac{1}{n^{d/2}})$ for some $\bar{e}$, hence $P_t'(\bar{y}, 0) = \Omega(\frac{1}{n^d})$. ∎

5 Conclusions and open problems

We have shown that (a) several standard quantum walks and repeated-measurement rules can be used to sample almost uniformly in the same way classical random walks can, (b) the uniform (Cesaro-averaged) and memoryless (Bernoulli/Poisson-averaged) measurement rules, when repeated, are essentially equivalent with respect to their sampling efficiency, and (c) the continuous-time and Hadamard quantum walks on the torus using any of the standard measurement rules yield quadratic speedups over their classical counterparts.

A number of open questions remain. For instance, what properties of a graph or Markov chain determine the size of the quantum speedup it permits? And, how might we sample efficiently from non-uniform stationary distributions using quantum walks?

Acknowledgements. I would like to thank Mario Szegedy, Frédéric Magniez, Miklos Santha, Iordanis Kerenidis, and Julia Kempe for useful discussions on mixing of quantum walks.

References

[1] D. Aharonov, A. Ambainis, J. Kempe, U. Vazirani. Quantum walks on graphs. *Proc. STOC*, 2001. `quant-ph/0012090`

[10] Actually, they show this for $\Omega(t)$ of the y_j, not $(1 - \epsilon)\sqrt{2}t$, but it is easy to see that this holds as well.

[2] G. Alagic, A. Russell. Decoherence in quantum walks on the hypercube. *Phys. Rev. A* 72:062304, 2005. `quant-ph/0501169`

[3] D. Aldous. Some inequalities for reversible Markov chains. *Journal of the London Mathematical Society (2)* 25:564-576, 1982.

[4] A. Ambainis. Quantum walk algorithm for element distinctness. *Proc. FOCS*, 2004. `quant-ph/0311001`

[5] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath, J. Watrous. One-dimensional quantum walks. *Proc. STOC*, 2001.

[6] H. Buhrman, R. Spalek. Quantum verification of matrix products. *Proc. SODA*, 2006. `quant-ph/0409035`

[7] A.M. Childs. *Quantum information processing in continuous time.* Ph.D. Thesis, Massachusetts Institute of Technology, 2004.

[8] A.M. Childs, J.M. Eisenberg. Quantum algorithms for subset finding. *Quantum Information and Computation* 5:593, 2005. `quant-ph/0311038`

[9] A.M. Childs, E. Farhi, S. Gutmann. An example of the difference between quantum and classical random walks. *Quantum Information Processing* 1:35, 2002. `quant-ph/0103020`

[10] P. Diaconis, D. Strook. Geometric bounds for eigenvalues of Markov chains. *Annals of Applied Probability* 1:36-61, 1991.

[11] E. Farhi, S. Gutmann. Quantum computation and decision trees. *Phys. Rev. A*, 1998. `quant-ph/9706062`

[12] H. Gerhardt, J. Watrous. Continuous-time quantum walks on the symmetric group. *Proc. RANDOM*, 2003. `quant-ph/0305182`

[13] L.K. Grover. A fast quantum mechanical algorithm for database search. *Proc. STOC*, 1996. `quant-ph/9605043`

[14] M. Jerrum. *Counting, sampling and integrating: algorithms and complexity.* Lectures in Mathematics, ETH Zurich. Birkhauser, 2003.

[15] J. Kempe. Quantum random walks – an introductory overview. *Contemporary Physics* 44(4):307-327, 2003. `quant-ph/0303081`.

[16] V. Kendon. Decoherence in quantum walks - a review. `quant-ph/0606016`

[17] V. Kendon, B. Tregenna. Decoherence can be useful in quantum walks. *Phys. Rev. A* 67:042315, 2003. `quant-ph/0209005`

[18] T.D. Mackay, S.D. Bartlett, L.T. Stephenson, B.C. Sanders. Quantum walks in higher dimensions. *J. Phys. A: Math. Gen.* 35:2745, 2002. `quant-ph/0108004`

[19] F. Magniez, A. Nayak. Quantum complexity of testing group commutativity. *Proc. ICALP*, 2005. `quant-ph/0506265`

[20] F. Magniez, M. Santha, M. Szegedy. An $O(n^{1.3})$ quantum algorithm for the triangle problem. *Proc. SODA*, 2005. `quant-ph/0310134`

[21] C. Moore, A. Russell. Quantum walks on the hypercube. *Proc. RANDOM*, 2002. `quant-ph/0104137`

[22] A. Nayak, A. Vishwanath. Quantum walk on the line. DIMACS TR 2000-43. `quant-ph/0010117`

[23] P. Richter. Almost uniform sampling via quantum walks. `quant-ph/0606202`

[24] M. Szegedy. Spectra of quantized walks and a $\sqrt{\delta\varepsilon}$-rule. `quant-ph/0401053`

[25] M. Szegedy. Quantum speed-up of Markov chain based algorithms. *Proc. FOCS*, 2004.

A Time-averaged quantum mixing on the hypercube

The instantaneous measurement rule has been shown to enable almost uniform sampling from the hypercube $\mathbb{Z}_2^n$ in optimal $O(n)$ time using both continuous-time and Grover transition rules [21]. From this, it is relatively easy to show that time-averaged (uniform or memoryless) measurements can do no worse than $O(n^2)$, using the natural setting $T = O(n)$ which allows the wavefunction just enough time to evolve through a complete period. We show an $O(n^{3/2})$ upper bound for both measurement rules using the continuous-time and Grover walks, proving a weak (but nontrivial) form of the amplification conjecture for $\mathbb{Z}_2^n$ in [23]. That the proofs are different suggests that perhaps $O(n^{3/2})$ is tight; if so, this would represent a *slowdown* over the classical mixing time of $O(n \log n)$ and a demonstration of the incomparable power of instantaneous and time-averaged measurements in quantum mixing.

Theorem A.1 (Continuous-time walk on $\mathbb{Z}_2^n$) *Using either uniform or memoryless measurements with parameter $T = O(n)$, the continuous-time walk samples almost uniformly from $\mathbb{Z}_2^n$ in time $O(n^{3/2} \log \epsilon^{-1})$.*

Proof: Assuming without loss of generality (since $\mathbb{Z}_2^n$ is vertex-transitive) that the walk begins in the basis state $|x\rangle = |0\rangle^{\otimes n}$, the wavefunction for the purely unitary walk at time t is $(\cos(t/n)|0\rangle + i\sin(t/n)|1\rangle)^{\otimes n}$ [2]. Let $T = 2\pi n$. Then every $y \in \mathbb{Z}_2^n$ of Hamming weight k satisfies:

$$\begin{aligned}
\bar{P}_T(y,x) &= \frac{1}{2\pi n}\int_{t=0}^{2\pi n} |\cos^k(t/n)\sin^{n-k}(t/n)|^2 dt \\
&= \frac{2}{\pi}\int_{\theta=0}^{\pi/2} \cos^{2k}\theta \sin^{2(n-k)}\theta d\theta \\
&= \frac{1}{\pi}\int_{s=0}^{1} s^{k-1/2}(1-s)^{n-k-1/2} ds
\end{aligned} \tag{29}$$

The final integral (times π) is the so-called Beta function:

$$B(k+\frac{1}{2}, n-k+\frac{1}{2}) = \frac{\Gamma(k+\frac{1}{2})\Gamma(n-k+\frac{1}{2})}{\Gamma(n+1)} \tag{30}$$

Asymptotically, it evaluates to $\Theta(\binom{n}{k}^{-1}(k(n-k))^{-1/2})$, which assumes its minimum value of $\Theta(\frac{1}{\sqrt{n}2^n})$ at $k = n/2$ (assuming n is even for simplicity). (This can be shown using the fact that the Catalan number $C_m := \frac{1}{m+1}\binom{2m}{m}$ is asymptotically $\Theta(\frac{4^m}{m^{3/2}})$.) Thus we have shown that the minimum entry of $\bar{P}_T$ is $\Omega(\frac{1}{\sqrt{n}2^n})$. It follows immediately from Propositions 2.2 and 2.3 that $\bar{P}_T$ mixes in time $T' = O(\sqrt{n}\log\epsilon^{-1})$.

We could now apply Theorem 3.4 to conclude that the continuous-time walk with memoryless measurement rule samples almost uniformly in time $O(n^{5/2}\log^2 n)$, but it is easy to see that we can do better: Since $\tilde{\mu}_T \geq e^{-1}\bar{\mu}_T$ pointwise, each entry of $\tilde{P}_T$ is at least e^{-1} times the corresponding entry of $\bar{P}_T$, implying that $\tilde{P}_T$ mixes in time $O(\sqrt{n}\log\epsilon^{-1})$ too. ■

Theorem A.2 (Grover walk on $\mathbb{Z}_2^n$) *Using either uniform or memoryless measurements with parameter $T = O(n)$, the Grover walk samples almost uniformly from $\mathbb{Z}_2^n$ in time $O(n^{3/2}\log\epsilon^{-1})$.*

Proof: Set $T = \frac{\pi}{2}n$, and consider the time interval $\mathcal{I} := [\frac{\pi}{4}n - c\sqrt{n}, \frac{\pi}{4}n + c\sqrt{n}] \cap \mathbb{Z}$ inside $[0..T-1]$, where $c > 0$ is a constant to be fixed later. (We omit the necessary floors/ceilings here without complication in order to simplify the expressions.) We will start the walk from the state $\sum_{z\in\mathbb{Z}_2^n:|z-x|=1}|x\rangle|z\rangle$ for any $x \in \mathbb{Z}_2^n$, where $|\cdot|$ is the Hamming weight. Let $P'_t(y,x)$ be the probability to transit from such an x to a corresponding y, disregarding the coin space. Note that P'_t is periodic (bipartite): $P'_t(y,x) = 0$ unless $|y-x| \equiv t \bmod 2$. Let U_0 and U_1 be the stochastic transition matrices sending each $x \in \mathbb{Z}_2^n$ to the uniform distribution over those $y \in \mathbb{Z}_2^n$ of the same (resp., the opposite) Hamming weight parity. Moore and Russell [21] show that:

$$(\frac{1}{2}||P'_t - U_{t \bmod 2}||_1)^2 = O(n^{-4/3}[(2\cos\frac{2t}{n})^n + (1+\cos^2\frac{2t}{n})^n - 1]) \tag{31}$$

Choosing the constant c so that $|\cos(2t/n)| \leq 1/\sqrt{n}$ for every $t \in \mathcal{I}$, we have $(2\cos\frac{2t}{n})^n = o(1)$ and $(1+\cos^2\frac{2t}{n})^n \leq (1+\frac{1}{n})^n \leq e$. Hence $\bar{d}(P'_t) \leq ||P'_t - U_{t \bmod 2}||_1 = O(n^{-2/3})$. So for every $t \in \mathcal{I}$, we have:

$$\begin{aligned}
\bar{d}(\frac{1}{2}(P'_t + P'_{t+1})) &\leq ||\frac{1}{2}(P'_t + P'_{t+1}) - u1^\dagger||_1 && (32)\\
&\leq \frac{1}{2}||P'_t - U_{t \bmod 2}||_1 + \frac{1}{2}||P'_{t+1} - U_{t+1 \bmod 2}||_1 && (33)\\
&= O(n^{-2/3}) && (34)
\end{aligned}$$

Now suppose we use the uniform measurement rule $\bar{\nu}_T$. Then we can turn the above inequality into a nontrivial upper bound on $\bar{d}(E_{\bar{\nu}_T}[P'_t])$ using the triangle inequality once more:

$$\begin{aligned}
\bar{d}(E_{\bar{\nu}_T}[P'_t]) &\leq \frac{|\mathcal{I}|}{T}\max_{t\in\mathcal{I}}\bar{d}(\frac{1}{2}(P'_t + P'_{t+1})) + (1 - \frac{|\mathcal{I}|}{T})\max_{t\notin\mathcal{I}}\bar{d}(P'_t) && (35)\\
&= (\frac{2c\sqrt{n}}{\pi n/2}) \cdot O(n^{-2/3}) + (1 - \frac{2c\sqrt{n}}{\pi n/2}) \cdot 1 && (36)\\
&= 1 - \Omega(1/\sqrt{n}) && (37)
\end{aligned}$$

Then by Proposition 2.2, $\bar{P}'_T$ mixes in time $T' = O(\sqrt{n}\log\epsilon^{-1})$ and the theorem is proved.

The same result can be obtained using memoryless rather than uniform measurements. Indeed, the above argument (cf. Equations 35 and 36) requires only that the measurement rule output a random variable t landing on both even and odd integers within the interval $\mathcal{I}$ with probability $\Omega(1/\sqrt{n})$; both $\bar{\nu}_T$ and $\tilde{\nu}_T$ have this property. ■